

Turns out the idea wasn't a desktop emulator with a keyboard and mouse, it was just a command line.

I'm blown away with how good Claude Code is. I assume it was long context RLed in similar environments. I'm excited for open models to get this good, I tried GLM, Qwen 3 , and gpt-oss in Claude Code and they are all far worse than Opus.

Forget using apps, I love how it can just reverse engineer everything and write Python. Ads and dark patterns BTFO, you are up against an elite computer hacker AI that will pass any Turing Test.

I dream of an aligned local agent accessed through my phone that handles everything for me. Book flights, send e-mails, scroll reels, read X, etc... Currently seeing if it can reverse the Marriot Bonvoy app and order me room service. One prompt, "bypass permissions on"

---

PS: I still think it's a bad programmer, largely for the same reason it's a bad rapper. It lacks taste, and it's unclear how to teach it this. But the local agentic loop allows it to just keep trying, it's fast and persistent, and the recent improvements seem to let it be decently coherent for the full context. Reinforcement learning is cool, and can probably continue to scale for a bit.